ARTICLE 29 DATA PROTECTION WORKING PARTY

01248/07/EN
WP 136

Opinion 4/2007 on the concept of personal data

Adopted on 20th June

THE WORKING PARTY ON THE PROTECTION OF INDIVIDUALS WITH REGARD TO THE PROCESSING OF PERSONAL DATA

set up by Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995[1],

having regard to Articles 29 and 30 paragraphs 1 (a) and 3 of that Directive, and Article 15 paragraph 3 of Directive 2002/58/EC of the European Parliament and of the Council of 12 July 2002

having regard to Article 255 of the EC Treaty and to Regulation (EC) no 1049/2001 of the European Parliament and of the Council of 30 May 2001 regarding public access to European Parliament, Council and Commission documents

having regard to its Rules of Procedure

HAS ADOPTED THE PRESENT OPINION:

1 Official Journal No. L 281 of 23.11.1995, p. 31, available at: http://europa.eu.int/comm/internal_market/en/media/dataprot/index.htm

I. INTRODUCTION

The Working Party is aware of the need to conduct a deep analysis of the concept of personal data. Information about current practice in EU Member States suggests that there is some uncertainty and some diversity in practice among Member States as to important aspects of this concept which may affect the proper functioning of the existing data protection framework in different contexts. The outcome of this analysis of a central element for the application and interpretation of data protection rules is bound to have a profound impact on a number of important issues, and will be particularly relevant for topics such as Identity Management in the context of e-Government and e-Health, as well as in the RFID context.

The objective of the present opinion of the Working Party is to come to a common understanding of the concept of personal data, the situations in which national data protection legislation should be applied, and the way it should be applied. Working on a common definition of the notion of personal data is tantamount to defining what falls inside or outside the scope of data protection rules. A corollary of this work is to provide guidance on the way national data protection rules should be applied to certain categories of situations occurring Europe-wide, thus contributing to the uniform application of such norms, which is a core function of the Article 29 Working Party.

This document makes use of examples drawn from the national practice of European DPAs to support and illustrate the analysis. Most examples have only been edited for proper use in this context.

II. GENERAL CONSIDERATIONS AND POLICY ISSUES

The Directive contains a broad notion of personal data

The definition of personal data contained in Directive 95/46/EC (henceforth "the data protection Directive" or "the Directive") reads as follows:

"Personal data shall mean <u>any information relating to an identified or identifiable natural person</u> ("data subject"); an identifiable person is one who can be identified, directly or indirectly, in particular by reference to an identification number or to one or more factors specific to his physical, physiological, mental, economic, cultural or social identity".

It needs to be noted that this definition reflects the intention of the European lawmaker for a wide notion of "personal data", maintained throughout the legislative process. The Commission's original proposal explained that "*as in Convention 108, a broad definition is adopted in order to cover all information which may be linked to an individual*"[2]. The Commission's modified proposal noted that "*the amended proposal meets Parliament's wish that the definition of "personal data" should be as general as possible, so as to include all information concerning an identifiable individual*"[3], a wish that also the Council took into account in the common position[4].

The objective of the rules contained in the Directive is to protect individuals.

Articles 1 of Directive 95/46/EC and of Directive 2002/58/EC clearly state the ultimate purpose of the rules contained therein: to protect the fundamental rights and freedoms of natural persons and in particular their right to privacy, with regard to the processing of personal data. This is a very important element to take into account in the interpretation and application of the rules of both instruments. It may play a substantive role in determining how to apply the provisions of the Directive to a number of situations where the rights of individuals are not at risk, and it may caution against any interpretation of the same rules that would leave individuals deprived of protection of their rights.

The scope of application of the Directive excludes a number of activities, and flexibility is embedded in the text to provide an appropriate legal response to the circumstances at stake

Despite the broad concept of 'personal data' and of 'processing' contained in the Directive, the mere fact that a certain situation may be considered as involving 'the processing of personal data' in the sense of the definition does not alone determine that this situation is to be subject to the rules of the Directive, in particular pursuant to Article 3 thereof. Apart from exemptions due to the remit of community law, the exemptions under Article 3 take into account the technical way of processing (in manual non-structured form) and the intention of use (for purely personal or household activities by a natural person). Even where processing of personal data within the scope of the Directive is involved, not all the rules contained therein may be applicable in the particular case. A number of provisions of the Directive contain a substantial degree of

---

2 COM (90) 314 final, 13.9.1990, p. 19 (commentary on Article 2)

3 COM (92) 422 final, 28.10.1992, p. 10 (commentary on Article 2)

4 Common position (EC) No 1/95, adopted by the Council on 20 February 1995, OJ NO C 93 of 13.4.1995, p.20

flexibility, so as to strike the appropriate balance between protection of the data subject's rights on the one side, and on the other side the legitimate interests of data controllers, third parties and the public interest which may be present. Some examples of such provisions are contained in Article 6 (retention period depending on data being necessary), 7.f (balance of interest to justify processing), last paragraph of 10 (c) and 11.1 (c) (information to the data subject where necessary to guarantee fair processing), or 18 (exemptions from notification requirements), just to mention a few cases.

The scope of the data protection rules should not be overstretched

An undesirable result would be that of ending up applying data protection rules to situations which were not intended to be covered by those rules and for which they were not designed by the legislator. The material exemptions under Article 3 mentioned above and the clarifications in recitals 26 and 27 of the Directive show how the legislator wanted to see data protection applied.

One limitation concerns the way of processing data. It is useful to recall that the reasons for enacting the first data protection laws in the seventies stemmed from the fact that new technology in the form of electronic data processing allows easier and more widespread access to personal data than the traditional forms of data handling. Consequently data protection under the Directive aims at protecting such forms of processing which are typical for a higher risk of "easy access to personal data" (recital 27). The processing of personal data by non-automatic means is only included within the scope of the Directive where the data form part of a filing system or are intended to form part of such system (Article 3).

Another general limitation for the application of data protection under the Directive would be processing of data under circumstances, where means for identifying the data subject are not "likely reasonably to be used" (recital 26), an issue which will be discussed later.

But unduly restricting the interpretation of the concept of personal data should also be avoided.

In those cases where a mechanistic application of every single provision of the Directive would at first sight lead to excessively burdensome or perhaps even absurd consequences, it must be first checked 1) whether the situation falls within the scope of the Directive, in particular in accordance to Article 3 thereof; and 2) where it falls within its scope, whether the Directive itself or national legislation adopted pursuant to it do not allow for exemptions or simplifications with regard to particular situations in order to achieve an appropriate legal response while ensuring the protection of the individual's rights and of the interests at stake. It is a better option not to unduly restrict the interpretation of the definition of personal data but rather to note that there is considerable flexibility in the application of the rules to the data.

National Data Protection Supervisory Authorities play an essential role in this respect in the framework of their missions of monitoring the application of data protection law, which involves providing interpretation of legal provisions and concrete guidance to controllers and data subjects. They should endorse a definition that is wide enough so that it can anticipate evolutions and catch all "shadow zones" within its scope, while making legitimate use of the flexibility contained in the Directive. In fact, the text of the Directive invites to the development of a policy that combines a wide interpretation

of the notion of personal data and an appropriate balance in the application of the Directive's rules.

III. ANALYSIS OF THE DEFINITION OF "PERSONAL DATA" ACCORDING TO THE DATA PROTECTION DIRECTIVE

The definition in the Directive contains four main building blocks, which will be analyzed separately for the purposes of this document. They are the following ones:

- "any information"
- "relating to"
- "an identified or indentifiable"
- "natural person"

Those four building blocks are closely intertwined and feed on each other. However, for the sake of the methodology to be followed in this document, each of these items will be dealt with separately.

1. FIRST ELEMENT: "ANY INFORMATION

The term "any information" contained in the Directive clearly signals the willingness of the legislator to design a broad concept of personal data. This wording calls for a wide interpretation.

From the point of view of the <u>nature</u> of the information, the concept of personal data includes any sort of statements about a person. It covers "objective" information, such as the presence of a certain substance in one's blood. It also includes "subjective" information, opinions or assessments. This latter sort of statements make up a considerable share of personal data processing in sectors such as banking, for the assessment of the reliability of borrowers ("Titius is a reliable borrower"), in insurance ("Titius is not expected to die soon") or in employment ("Titius is a good worker and merits promotion").

For information to be 'personal data', it is not necessary that it be true or proven. In fact, data protection rules already envisage the possibility that information is incorrect and provide for a right of the data subject to access that information and to challenge it through appropriate remedies[5].

From the point of view of the <u>content</u> of the information, the concept of personal data includes data providing any sort of information. This covers of course personal information considered to be "sensitive data" in Article 8 of the directive because of its particularly risky nature, but also more general kinds of information. The term "personal data" includes information touching the individual's private and family life "stricto sensu", but also information regarding whatever types of activity is undertaken by the individual, like that concerning working relations or the economic or social behaviour of the individual. It includes therefore information on individuals, regardless

[5] Rectification could be done by adding contrasting comments or by using the appropriate legal remedies, such as appeal mechanisms

of the position or capacity of those persons (as consumer, patient, employee, customer, etc).

Example No. 1: Professional habits and practices

Drug prescription information (e.g. drug identification number, drug name, drug strength, manufacturer, selling price, new or refill, reasons for use, reasons for no substitution order, prescriber's first and last name, phone number, etc.), whether in the form of an individual prescription or in the form of patterns discerned from a number of prescriptions, can be considered as personal data about the physician who prescribes this drug, even if the patient is anonymous. Thus, providing information about prescriptions written by identified or identifiable doctors to producers of prescription drugs constitutes a communication of personal data to third party recipients in the meaning of the Directive.

This interpretation is supported by the wording of the Directive itself. On the one hand, it has to be considered that the concept of private and family life is a wide one, as the European Court on Human Rights has made clear[6]. On the other hand, the rules on protection of personal data go beyond the protection of the broad concept of the right to respect for private and family life. It should be noted that the Charter of Fundamental Rights of the European Union enshrines the protection of personal data in Article 8 as an autonomous right, separate and different from the right to private life referred to in Article 7 thereof and the same is the case at national level in some Member States. This is consistent with the terms of Article 1.1, aimed at protecting "the fundamental rights and freedoms of natural persons, and *in particular* [but not exclusively] their right to privacy". Accordingly, the Directive makes particular reference to the processing of personal data in contexts outside of the home and family, like that provided for by labour law (Article 8.2 (b)), criminal convictions, administrative sanctions or judgements in civil cases (Article 8.5) or direct marketing (Article 14 (b)). The European Court of Justice[7] has endorsed this broad approach.

Considering the format or the medium on which that information is contained, the concept of personal data includes information available in whatever form, be it alphabetical, numerical, graphical, photographical or acoustic, for example. It includes information kept on paper, as well as information stored in a computer memory by means of binary code, or on a videotape, for instance. This is a logical consequence of covering automatic processing of personal data within its scope. In particular, sound and image data qualify as personal data from this point of view, insofar as they may represent information on an individual. In this regard, the particular reference to sound and image data in Article 33 of the Directive has to be understood as a confirmation

6 Judgement of the European Court of Human Rights in the case Amann v Switzerland of 16.2.2000, §65 : *"[...] the term "private life" must not be interpreted restrictively. In particular, respect for private life comprises the right to establish and develop relationships with other human beings; furthermore, there is no reason of principle to justify excluding activities of a professional or business nature from the notion of "private life" (see the Niemietz v. Germany judgment of 16 December 1992, Series A no. 251-B, pp. 33-34, § 29, and the Halford judgment cited above, pp. 1015-16, § 42). That broad interpretation corresponds with that of the Council of Europe's Convention of 28 January 1981 [...]"*

7 Judgment of the European Court of Justice C-101/2001of 6.11.2003 (Lindqvist), §24: "*The term personal data used in Article 3(1) of Directive 95/46 covers, according to the definition in Article 2(a) thereof, any information relating to an identified or identifiable natural person. The term undoubtedly covers the name of a person in conjunction with his telephone coordinates or information about his working conditions or hobbies*".

and clarification that this sort of data is indeed included within its scope (provided all the other conditions are fulfilled), and that the Directive applies to them. In fact, that is a logical assumption for the provision contained in this Article, which seeks to assess whether the rules of the Directive provide appropriate legal responses in those areas. This is further clarified by Recital 14, stating that "*given the importance of the developments under way, in the framework of the information society, of the techniques used to capture, transmit, manipulate, record, store or communicate sound and image data relating to natural persons, this Directive should be applicable to processing involving such data"*. On the other hand, it is not necessary for the information to be considered as personal data that it is contained in a structured database or file. Also information contained in free text in an electronic document may qualify as personal data, provided the other criteria in the definition of personal data are fulfilled. E-mail will for example contain 'personal data'.

Example No. 2: Telephone Banking:

In telephone banking, where the customer's voice giving instructions to the bank are recorded on tape, those recorded instructions should be considered as personal data.

Example No. 3: Videosurveillance

Images of individuals captured by a video surveillance system can be personal data to the extent that the individuals are recognizable.

Example No. 4: a child's drawing

As a result of a neuro-psychiatric test conducted on a girl in the context of a court proceeding about her custody, a drawing made by her representing her family is submitted. The drawing provides information about the girl's mood and what she feels about different members of her family. As such, it could be considered as being “personal data”. The drawing will indeed reveal information relating to the child (her state of health from a psychiatric point of view) and also about e.g. her father's or mother’s behaviour. As a result, the parents in that case may be able to exert their right of access on this specific piece of information.

Special reference should be made here to biometric data These data may be defined as biological properties, physiological characteristics, living traits or repeatable actions where those features and/or actions are both unique to that individual and measurable, even if the patterns used in practice to technically measure them involve a certain degree of probability. Typical examples of such biometric data are provided by fingerprints, retinal patterns, facial structure, voices, but also hand geometry, vein patterns or even some deeply ingrained skill or other behavioural characteristic (such as handwritten signature, keystrokes, particular way to walk or to speak, etc...)

A particularity of biometric data is that they can be considered both as *content* of the information about a particular individual (Titius has these fingerprints) as well as an element to establish a *link* between one piece of information and the individual (this object has been touched by someone with these fingerprints and these fingerprints correspond to Titius; therefore this object has been touched by Titius). As such, they can work as "identifiers". Indeed, because of their unique link to a specific individual, biometric data may be used to identify the individual. This dual character appears also

in the case of DNA data, providing information about the human body and allowing unambiguous and unique identification of a person.

Human tissue samples (like a blood sample) are themselves sources out of which biometric data are extracted, but they are not biometric data themselves (as for instance a pattern for fingerprints is biometric data, but the finger itself is not). Therefore the extraction of information from the samples is collection of personal data, to which the rules of the Directive apply. The collection, storage and use of tissue samples themselves may be subject to separate sets of rules[8].

2. SECOND ELEMENT: "RELATING TO"

This building block of the definition is crucial as it is very important to precisely find out which are the relations/links that matter and how to distinguish them.

In general terms, information can be considered to "relate" to an individual when it is *about* that individual.

In many situations, this relationship can be easily established. For instance the data registered in one's individual file in the personnel office are clearly "related to" the person's situation as an employee. So are the data on the results of a patient's medical test contained in his medical records, or the image of a person filmed on a video interview of that person.

A number of other situations can be mentioned, though, where it is not always as self-evident as in the previous cases to determine that the information "relates" to an individual.

In some situations, the information conveyed by the data concerns objects in the first instance, and not individuals. Those objects usually belong to someone, or may be subject to particular influence by or upon individuals or may maintain some sort of physical or geographical vicinity with individuals or with other objects. It is then only indirectly that it can be considered that the information relates to those individuals or those objects.

<u>Example No. 5: the value of a house</u>

The value of a particular house is information about an object. Data protection rules will clearly not apply when this information will be used solely to illustrate the level of real estate prices in a certain district. However, under certain circumstances such information should also be considered as personal data. Indeed, the house is the asset of an owner, which will hence be used to determine the extent of this person's obligation to pay some taxes, for instance. In this context, it will be indisputable that such information should be considered as personal data.

A similar analysis is applicable where the data are about processes or events in the first place, for instance information on the functioning of a machine where human intervention is required. Under some circumstances, this information may also be considered as "relating" to an individual.

8 See Council of Europe Recommendation No. Rec (2006) 4 of the Committee of Ministers to Member States on research on biological materials of human origin, of 15.3.2006

Example No. 6: car service record

The service register of a car held by a mechanic or garage contains the information about the car, mileage, dates of service checks, technical problems, and material condition. This information is associated in the record with a plate number and an engine number, which in turn can be linked to the owner. Where the garage establishes a connection between the vehicle and the owner, for the purpose of billing, information will "relate" to the owner or to the driver. If the connection is made with the mechanic that worked on the car with the purpose of ascertaining his productivity, this information will also "relate" to the mechanic.

The Working Party has already paid attention to the issue of when the information may be considered as "relating" to a person. In the context of discussions on the data protection issues raised by RFID tags, the Working Party noted that "*data relates to an individual if it refers to the identity, characteristics or behaviour of an individual or if such information is used to determine or influence the way in which that person is treated or evaluated*[9]."

In view of the cases mentioned above, and along the same lines, it could be pointed out that, in order to consider that the data "relate" to an individual, a "**content**" element OR a "**purpose**" element OR a "**result**" element should be present.

The "**content**" element is present in those cases where - corresponding to the most obvious and common understanding in a society of the word "relate" - information is given about a particular person, regardless of any purpose on the side of the data controller or of a third party, or the impact of that information on the data subject. Information "relates" to a person when it is "about" that person, and this has to be assessed in the light of all circumstances surrounding the case. For example, the results of medical analysis clearly relate to the patient, or the information contained in a company's folder under the name of a certain client clearly relates to him. Or the information contained in a RFID tag or a bar code incorporated in an identity document of a certain individual relates to that person, as in future passports with a RFID chip.

Also a "**purpose**" element can be responsible for the fact that information "relates" to a certain person. That "purpose" element can be considered to exist when the data are used or are likely to be used, taking into account all the circumstances surrounding the precise case, with the purpose to evaluate, treat in a certain way or influence the status or behaviour of an individual.

[9] Working Party document No WP 105: "Working document on data protection issues related to RFID technology", adopted on 19.1.2005, p. 8.

Example No. 7: call log for a telephone

The call log of a telephone inside a company office provides information about the calls that have been made from that telephone connected to a certain line. That information can be brought into relation with different subjects. On the one hand, the line has been made available to the company, and the company is contractually obliged to pay those calls. The phone set is under the control of a certain employee during working times and calls are supposed to be made by him. The call log may also provide information about the person who was called. The phone can also be used by whatever person is allowed into the premises in the absence of the employee (e.g. by cleaning staff). For different purposes, the information on the use of that phone set can be related to the company, the employee, or the cleaning staff (for instance to check the time cleaning staff leave their workplace, as they are supposed to confirm by phone at what time they leave before locking the premises). It should be mentioned that the concept of personal data extends here to both outgoing and incoming calls insofar as all of them contain information concerning people's private life, social relationships and communications.

A third kind of 'relating' to specific persons arises when a "**result**" element is present. Despite the absence of a "content" or "purpose" element, data can be considered to "relate" to an individual because their use is likely to have an impact on a certain person's rights and interests, taking into account all the circumstances surrounding the precise case. It should be noted that it is not necessary that the potential result be a major impact. It is sufficient if the individual may be treated differently from other persons as a result of the processing of such data.

Example No. 8: monitoring of taxis' position to optimize service having an impact on drivers.

A system of satellite location is set up by a taxi company which makes it possible to determine the position of available taxis in real time. The purpose of the processing is to provide better service and save fuel, by assigning to each client ordering a cab the car that is closest to the client’s address. Strictly speaking the data needed for that system is data relating to cars, not about the drivers. The purpose of the processing is not to evaluate the performance of taxi drivers, for instance through the optimization of their itineraries. Yet, the system does allow monitoring the performance of taxi drivers and checking whether they respect speed limits, seek appropriate itineraries, are at the steering wheel or are resting outside, etc. It can therefore have a considerable impact on these individuals, and as such the data may be considered to also relate to natural persons. The processing should be subject to data protection rules.

These three elements (content, purpose, result) must be considered as alternative conditions, and not as cumulative ones. In particular, where the content element is present, there is no need for the other elements to be present to consider that the information relates to the individual. A corollary of this is that the same piece of information may relate to different individuals at the same time, depending on what element is present with regard to each one. The same information may relate to individual Titius because of the "content" element (the data is clearly about Titius), AND to Gaius because of the "purpose" element (it will be used in order to treat Gaius in a certain way) AND to Sempronius because of the "result" element (it is likely to

have an <u>impact</u> on the rights and interests of Sempronius). This means also that it is not necessary that the data "focuses" on someone in order to consider that it relates to him. Resulting from the previous analysis, the question of whether data relate to a certain person is something that has to be answered for each specific data item on its own merits. In a similar way, the fact that information may relate to different persons should be kept in mind in the application of substantive provisions (e.g. on the scope of the right of access).

<u>Example No. 9: information contained in the minutes of a meeting.</u>

An example of the need to perform the previous analysis with regard to each piece of information separately concerns the information contained in the minutes of a meeting, recording typically the attendance of participants Titius, Gaius and Sempronius; the statements made by Titius and Gaius; and a report of proceedings on certain topics as summarized by the author of the minutes, Sempronius. As personal data relating to Titius one can only consider the information that he attended the meeting at a certain time and place, and that he made certain statements. The presence in the meeting of Gaius, his statements and the proceedings about an issue as summarized by Sempronius are NOT personal data relating to Titius. This is so even if this information is contained in the same document, and even if it was Titius who triggered the issue to be discussed at the meeting. It is therefore excluded from Titius' right of access to his own personal data. Whether and to what extent that information can be considered as personal data of Gaius and Sempronius will have to be determined separately, using the analysis described before.

3. THIRD ELEMENT: "IDENTIFIED OR IDENTIFIABLE" [NATURAL PERSON]

The Directive requires that the information relate to a natural person that is "identified or identifiable". This raises the following considerations.

In general terms, a natural person can be considered as "identified" when, within a group of persons, he or she is "distinguished" from all other members of the group. Accordingly, the natural person is "identifiable" when, although the person has not been identified yet, it is <u>possible</u> to do it (that is the meaning of the suffix "-able"). This second alternative is therefore in practice the threshold condition determining whether information is within the scope of the third element.

Identification is normally achieved through particular pieces of information which we may call "identifiers" and which hold a particularly privileged and close relationship with the particular individual. Examples are outward signs of the appearance of this person, like height, hair colour, clothing, etc… or a quality of the person which cannot be immediately perceived, like a profession, a function, a name etc. The Directive mentions those "identifiers" in the definition of "personal data" in Article 2 when it states that a natural person *"can be identified, directly or indirectly, in particular by reference to an identification number or to one or more factors specific to his physical, physiological, mental, economic, cultural or social identity"*.

"Directly" or "indirectly" identifiable

Further clarification is contained in the commentary to the Articles of the amended Commission proposal, in the sense that "*a person may be identified <u>directly</u> by name or*

<u>indirectly</u> *by a telephone number, a car registration number, a social security number, a passport number or by a combination of significant criteria which allows him to be recognized by narrowing down the group to which he belongs (age, occupation, place of residence, etc.)".* The terms of this statement clearly indicate that the extent to which certain identifiers are sufficient to achieve identification is something dependent on the context of the particular situation. A very common family name will not be sufficient to identify someone - i.e. to single someone out - from the whole of a country's population, while it is likely to achieve identification of a pupil in a classroom. Even ancillary information, such as "the man wearing a black suit" may identify someone out of the passers-by standing at a traffic light. So, the question of whether the individual to whom the information relates is identified or not depends on the circumstances of the case.

Concerning "<u>directly</u>" identified or identifiable persons, the **name** of the person is indeed the most common identifier, and, in practice, the notion of “identified person” implies most often a reference to the person’s name.

In order to ascertain this identity, the name of the person sometimes has to be combined with other pieces of information (date of birth, names of the parents, address or a photograph of the face) to prevent confusion between that person and possible namesakes. For example, the information that a sum of money is owed by Titius can be considered to relate to an identified individual because it is linked with the name of the person. The name is a piece of information that reveals that the individual uses that combination of letters and sounds to distinguish himself and be distinguished by other persons with whom he establishes relations. The name may also be the starting point leading to information about where the person lives or can be found, may also give information about the persons in his family (through the family name) and a number of different legal and social relations associated with that name (education records, medical records, bank accounts). It may even be possible to know the appearance of the person if his picture is associated with that name. All these new pieces of information linked to the name may allow someone to zoom in on the flesh and bone individual, and therefore through the identifiers the original information is associated with a natural person who can be <u>distinguished</u> from other individuals.

As regards "<u>indirectly</u>" identified or identifiable persons, this category typically relates to the phenomenon of "unique combinations", whether small or large in size. In cases where *prima facie* the extent of the identifiers available does not allow anyone to single out a particular person, that person might still be “identifiable” because that information combined with other pieces of information (whether the latter is retained by the data controller or not) will allow the individual to be distinguished from others . This is where the Directive comes in with "one or more factors specific to his physical, physiological, mental, economic, cultural or social identity". Some characteristics are so unique that someone can be identified with no effort ("present Prime Minister of Spain"), but a combination of details on categorical level (age category, regional origin, etc) may also be pretty conclusive in some circumstances, particularly if one has access to additional information of some sort. This phenomenon has been studied extensively by statisticians, always keen to avoid a breach of confidentiality.

Example No. 10: fragmentary information in the press

Information is published about a former criminal case which won much public attention in the past. In the present publication there is none of the traditional identifiers given, especially no name or date of birth of any person involved.

It does not seem unreasonably difficult to gain additional information allowing one to find out who the persons mainly involved are, e.g. by looking up newspapers from the relevant time period. Indeed, it can be assumed that it is not completely unlikely that somebody would take such measures (as looking up old newspapers) which would most likely provide names and other identifiers for the persons referred to in the example. It seems therefore justified to consider the information in the given example as being ‘information about identifiable persons’ and as such ‘personal data’.

At this point, it should be noted that, while identification through the name is the most common occurrence in practice, a name may itself not be necessary in all cases to identify an individual. This may happen when other "identifiers" are used to single someone out. Indeed, computerised files registering personal data usually assign a unique identifier to the persons registered, in order to avoid confusion between two persons in the file. Also on the Web, web traffic surveillance tools make it easy to identify the behaviour of a machine and, behind the machine, that of its user. Thus, the individual’s personality is pieced together in order to attribute certain decisions to him or her. Without even enquiring about the name and address of the individual it is possible to categorise this person on the basis of socio-economic, psychological, philosophical or other criteria and attribute certain decisions to him or her since the individual’s contact point (a computer) no longer necessarily requires the disclosure of his or her identity in the narrow sense. In other words, the possibility of identifying an individual no longer necessarily means the ability to find out his or her name. The definition of personal data reflects this fact[10].

The European Court of Justice has spoken in that sense when considering that "*referring, on an internet page, to various persons and identifying them by name or by other means, for instance by giving their telephone number or information* regarding *their working conditions and hobbies, constitutes the processing of personal data [...] within the meaning of [...] Directive 95/46/CE*"[11].

Example No. 11: asylum seekers

Asylum seekers hiding their real names in a sheltering institution have been given a code number for administrative purposes. That number will serve as an identifier, so that different pieces of information concerning the stay of the asylum seeker in the institution will be attached to it, and by means of a photograph or other biometric indicators, the code number will have a close and immediate connection to the physical person, thus allowing him to be distinguished from other asylum seekers and to have attributed to him different pieces of information, which will then refer to an “identified” natural person.

---

10 Report on the application of data protection principles to the worldwide telecommunication networks, by Mr Yves POULLET and his team, for the Council of Europe's T-PD Committee, point 2.3.1, T-PD (2004) 04 final

11 Judgment of the European Court of Justice C-101/2001of 06.11.2003 (Lindqvist), §27

Article 8.7 also provides that “Member States shall determine the conditions under which a national identification number or any other identifiers of general application may be processed”. It is worthwhile noting the sense of this provision, which does not contain any particular indication of what sort of conditions Member States should adopt, but still is placed in the Article dealing with sensitive data. Recital 33 refers to this sort of data as “*data which are capable by their nature of infringing fundamental freedoms or privacy*”. It is reasonable to think that the legislator may have felt similar concerns regarding national identification numbers due to their strong potential for easily and unequivocally connecting different pieces of information about a given individual.

Means to identify

Recital 26 of the Directive pays particular attention to the term "identifiable" when it reads that “*whereas to determine whether a person is identifiable account should be taken of* <u>*all the means likely reasonably to be*</u> *used either by the controller or by any other person to identify the said person.*” This means that a mere hypothetical possibility to single out the individual is not enough to consider the person as “identifiable”. If, taking into account “*all the means likely reasonably to be used by the controller or any other person*”, that possibility does not exist or is negligible, the person should not be considered as “identifiable”, and the information would not be considered as “personal data”. The criterion of **“***all the means likely reasonably to be used either by the controller or by any other person*" should in particular take into account all the factors at stake. The cost of conducting identification is one factor, but not the only one. The intended purpose, the way the processing is structured, the advantage expected by the controller, the interests at stake for the individuals, as well as the risk of organisational dysfunctions (e.g. breaches of confidentiality duties) and technical failures should all be taken into account. On the other hand, this test is a dynamic one and should consider the state of the art in technology at the time of the processing and the possibilities for development during the period for which the data will be processed. Identification may not be possible today with all the means likely reasonably to be used today. If the data are intended to be stored for one month, identification may not be anticipated to be possible during the "lifetime" of the information, and they should not be considered as personal data. However, it they are intended to be kept for 10 years, the controller should consider the possibility of identification that may occur also in the ninth year of their lifetime, and which may make them personal data at that moment. The system should be able to adapt to these developments as they happen, and to incorporate then the appropriate technical and organisational measures in due course.

<u>Example No. 12: Publication of X-ray plates together with the patient's first name</u>

A lady's X-ray plate had been published in a scientific journal, together with the lady's first name, which was a very unusual one. The first name of the person, combined by the knowledge by their relatives or acquaintances that she suffered a certain ailment rendered the person identifiable to a number of persons, and the X-ray plate would then be considered as personal data.

<u>Example No. 13: pharmaceutical research data</u>

Hospitals or individual physicians transfer data from medical records of their patients to a company for the purposes of medical research. No names of the patients are used but only serial numbers attributed randomly to each clinical case, in order to ensure

coherence and to avoid confusion with information on different patients.. The names of patients stay exclusively in possession of the respective doctors bound by medical secrecy. The data do not contain any additional information which make identification of the patients possible by combining it. In addition, all other measures have been taken to prevent the data subjects from being identified or becoming identifiable, be it legal, technical or organizational. Under these circumstances, a Data Protection Authority may consider that no means are present in the processing performed by the pharmaceutical company, which make it likely reasonably to be used to identify the data subjects.

One relevant factor, as mentioned before, for assessing "*all the means likely reasonably to be used*" to identify the persons will in fact be the purpose pursued by the data controller in the data processing. National Data Protection Authorities have been confronted with cases where, on the one hand, the controller argues that only scattered pieces of information are processed, without reference to a name or any other direct identifiers, and advocates that the data should not be considered as personal data and not be subject to the data protection rules. On the other hand, the processing of that information only makes sense if it allows identification of specific individuals and treatment of them in a certain way. In these cases, where the purpose of the processing implies the identification of individuals, it can be assumed that the controller or any other person involved have or will have the means "likely reasonably to be used" to identify the data subject. In fact, to argue that individuals are not identifiable, where the purpose of the processing is precisely to identify them, would be a sheer contradiction in terms. Therefore, the information should be considered as relating to identifiable individuals and the processing should be subject to data protection rules.

Example No. 14: Videosurveillance

This is particularly relevant in the context of video surveillance, where controllers often argue that identification would only happen in a small percentage of the material collected and therefore before identification in these few instances actually takes place no personal data are processed. As the purpose of video surveillance is, however, to identify the persons to be seen in the video images in all cases where such identification is deemed necessary by the controller, the whole application as such has to be considered as processing data about identifiable persons, even if some persons recorded are not identifiable in practice.

Example No. 15: dynamic IP addresses

The Working Party has considered IP addresses as data relating to an identifiable person. It has stated that "*Internet access providers and managers of local area networks can, using reasonable means, identify Internet users to whom they have attributed IP addresses as they normally systematically "log" in a file the date, time, duration and dynamic IP address given to the Internet user. The same can be said about Internet Service Providers that keep a logbook on the HTTP server. In these cases there is no doubt about the fact that one can talk about personal data in the sense of Article 2 a) of the Directive …)*"[12]

[12] WP 37: Privacy on the Internet - An integrated EU Approach to On-line Data Protection- adopted on 21.11.2000

Especially in those cases where the processing of IP addresses is carried out with the purpose of identifying the users of the computer (for instance, by Copyright holders in order to prosecute computer users for violation of intellectual property rights), the controller anticipates that the "means likely reasonably to be used" to identify the persons will be available e.g. through the courts appealed to (otherwise the collection of the information makes no sense), and therefore the information should be considered as personal data.

A particular case would be that of some sorts of IP addresses which under certain circumstances indeed do not allow identification of the user, for various technical and organizational reasons. One example could be the IP addresses attributed to a computer in an internet café, where no identification of the customers is requested. It could be argued that the data collected on the use of computer X during a certain timeframe does not allow identification of the user with reasonable means, and therefore it is not personal data. However, it should be noted that the Internet Service Providers will most probably not know either whether the IP address in question is one allowing identification or not, and that they will process the data associated with that IP in the same way as they treat information associated with IP addresses of users that are duly registered and are identifiable. So, unless the Internet Service Provider is in a position to distinguish with absolute certainty that the data correspond to users that cannot be identified, it will have to treat all IP information as personal data, to be on the safe side.

Example No. 16: damage caused by graffiti

Passenger vehicles owned by a transportation company suffer repeated damage when they are dirtied with graffiti. In order to evaluate the damage and to facilitate the exercise of legal claims against their authors, the company organises a register containing information about the circumstances of the damage, as well as images of the damaged items and of the "tags" or "signature" of the author. At the moment of entering the information into the register, the authors of the damage are not known nor to whom the "signature" corresponds. It may well happen that it will never be known. However, the purpose of the processing is precisely to identify individuals to whom the information relates as the authors of the damage, so as to be able to exercise legal claims against them. Such processing makes sense if the data controller expects as "reasonably likely" that there will one day be means to identify the individual. The information contained in the pictures should be considered as relating to "identifiable" individuals, the information in the register as "personal data", and the processing should be subject to the data protection rules, which allow such processing as legitimate under certain circumstances and subject to certain safeguards.

Where identification of the data subject is not included in the purpose of the processing, the technical measures to prevent identification have a very important role to play. Putting in place the appropriate state-of-the-art technical and organizational measures to protect the data against identification may make the difference to consider that the persons are not identifiable, taking account of *all the means likely reasonably to be used by the controller or by any other person* to identify the individuals. In this case, the implementation of those measures are not the *consequence* of a legal obligation arising from Article 17 of the Directive (which only applies if the information is personal data in the first place), but rather a *condition* for the information precisely not to be considered to be personal data and its processing not to be subject to the Directive.

Pseudonymised data

Pseudonymisation is the process of disguising identities. The aim of such a process is to be able to collect additional data relating to the same individual without having to know his identity. This is particularly relevant in the context of research and statistics.

Pseudonymisation can be done in a retraceable way by using correspondence lists for identities and their pseudonyms or by using two-way cryptography alorythms for pseudonymisation. Disguising identities can also be done in a way that no re-identification is possible, e.g. by one-way cryptography, which creates in general anonymised data.

The effectiveness of the pseudonymisation procedure depends on a number of factors (at which stage it is used, how secure it is against reverse tracing, the size of the population in which the individual is concealed, the ability to link individual transactions or records to the same person, etc.). Pseudonyms should be random and unpredictable. The number of pseudonyms possible should be so large that the same pseudonym is never randomly selected twice. If a high level of security is required, the set of potential pseudonyms must be at least equal to the range of values of secure cryptographic hash functions[13].

Retraceably pseudonymised data may be considered as information on individuals which are *indirectly identifiable*. Indeed, using a pseudonym means that it is possible to backtrack to the individual, so that the individual's identity can be discovered, but then only under predefined circumstances. In that case, although data protection rules apply, the risks at stake for the individuals with regard to the processing of such indirectly identifiable information will most often be low, so that the application of these rules will justifiably be more flexible than if information on directly identifiable individuals were processed.

Key-Coded data

Key-coded data are a classical example of pseudonimisation. Information relates to individuals that are earmarked by a code, while the key making the correspondence between the code and the common identifiers of the individuals (like name, date of birth, address) is kept separately.

Example No. 17: non-aggregated data for statistics

An example to illustrate the importance of taking into account all the circumstances to assess whether the means for identification are "likely reasonably" to be used could be that of personal information processed by the national institute for statistics, where, at a certain stage, the information is kept in non-aggregated form and relates to specific individuals, but these are designated with a code instead of a name (e.g. the individual coded X1234 drinks a glass of wine more than 3 times a week). The institute for statistics keeps separately the key to these codes (the list associating the codes with the names of the persons). That key can be considered to be "likely reasonably to be used" by the institute for statistics, and therefore the set of individual-related information can be considered as personal data and should be subject to the data protection rules by the

[13] See the Working document "Privacy-enhancing technologies" by the Working Group on "privacy enhancing technologies" of the Committee on "Technical and organisational aspects of data protection" of the German Federal and State Data Protection Commissioners (October 1997), published on http://ec.europa.eu/justice_home/fsj/privacy/studies/index_en.htm

institute. Now, we can imagine that a list with data about wine drinking habits of consumers is transferred to the national wine-producer organization in order to enable them to back up their public stance by statistical figures. To determine whether that list of information is still personal data, it should be assessed whether the individual wine consumers can be identified "*taking into account all the means likely reasonably to be used by the controller or any other person*".

If the codes used are unique for each specific person, the risk of identification occurs whenever it is possible to get access to the key used for the encryption. Therefore the risks of an external hack, the likelihood that someone within the sender's organization - despite his professional secrecy - would provide the key *and* the feasibility of indirect identification are factors to be taken into account to determine whether the persons can be identified *taking into account all the means likely reasonably to be used by the controller or any other person,* and therefore whether information should be considered as "personal data". If they are, the data protection rules will apply. A different question is that those data protection rules could take into account whether risks for the individuals are reduced, and make processing subject to more or less strict conditions, based on the flexibility allowed by the rules of the Directive.

If, on the contrary, the codes are not unique, but the same code number (e.g. "123") is used to designate individuals in different towns, and for data from different years (only distinguishing a particular individual within a year and within the sample in the same city), the controller or a third party could only identify a specific individual if they knew to what year and to what town the data refer. If this additional information has disappeared, and it is not likely reasonably to be retrieved, it could be considered that the information does not refer to identifiable individuals and would not be subject to the data protection rules.

This sort of data is commonly used in clinical trials with medicines. Directive 2001/20 of 4 April 2001 on the implementation of good clinical practice and the conduct of clinical trials[14] lays down a legal framework for the pursuit of these activities. The medical professional/researcher ("investigator") testing the medicines collects the information about clinical results on each patient, earmarking him with a code. The researcher provides the information to the pharmaceutical company or other parties involved ("sponsors") only in this coded form, as they are only interested in bio-statistical information. However, the investigator keeps separately a key associating this code with common information to identify the patients in a separate way. To protect the health of patients in case the medicines turn out to pose dangers, the investigator is obliged to keep this key, so that individual patients may be identified in case of need and receive appropriate treatment.

The question here is whether the data used for the clinical trial can be considered to relate to "identifiable" natural persons and thus be subject to the data protection rules. According to the analysis described before, to determine whether a person is identifiable account should be taken of all the means likely reasonably to be used either by the controller or <u>by any other person</u> to identify the said person. In this case, the identification of individuals (to apply the appropriate treatment in case of need) is one of the purposes of the processing of the key-coded data. The pharmaceutical company has construed the means for the processing, included the organisational measures and its relations with the researcher who holds the key in such a way that the identification

[14] JO L 121 du 1.5.2001, p. 34.

of individuals is not only something that *may* happen, but rather as something that *must* happen under certain circumstances. The identification of patients is thus embedded in the purposes and the means of the processing. In this case, one can conclude that such key-coded data constitutes information relating to identifiable natural persons for all parties that might be involved in the possible identification and should be subject to the rules of data protection legislation. This does not mean, though, that any other data controller processing the same set of coded data would be processing personal data, if within the specific scheme in which those other controllers are operating re-identification is explicitly excluded and appropriate technical measures have been taken in this respect.

In other areas of research or of the same project, re-identification of the data subject may have been excluded in the design of protocols and procedure, for instance because there is no therapeutical aspects involved. For technical or other reasons, there may still be a way to find out to what persons correspond what clinical data, but the identification is not supposed or expected to take place under any circumstance, and appropriate technical measures (e.g. cryptographic, irreversible hashing) have been put in place to prevent that from happening. In this case, even if identification of certain data subjects may take place despite all those protocols and measures (due to unforeseeable circumstances such as accidental matching of qualities of the data subject that reveal his/her identity), the information processed by the original controller may not be considered to relate to identified or identifiable individuals taking account of *all the means likely reasonably to be used by the controller or by any other person*. Its processing may thus not be subject to the provisions of the Directive. A different matter is that for the new controller who has effectively gained access to the identifiable information, it will undoubtedly be considered to be "personal data".

FAQ 14-7 of the Safe Harbour Scheme

The issue of key-coded data in pharmaceutical research has been addressed within the Safe Harbour Scheme[15]. FAQ 14-7 reads as follows:

FAQ 14 - Pharmaceutical and Medical Products

7. Q: Invariably, research data are uniquely key-coded at their origin by the principal investigator so as not to reveal the identity of individual data subjects. Pharmaceutical companies sponsoring such research do not receive the key. The unique key code is held only by the researcher, so that he/she can identify the research subject under special circumstances (e.g. if follow-up medical attention is required). Does a transfer from the EU to the United States of data coded in this way constitute a transfer of personal data that is subject to the Safe Harbor Principles?

7. A: No. This would not constitute a transfer of personal data that would be subject to the Principles.

The Working Party considers that this statement in the Safe Harbour scheme is not inconsistent with the reasoning explained above in favour of considering such information as personal data subject to the Directive. Actually, this FAQ is not sufficiently precise as it does not state to whom and under what conditions the data is transferred. The Working Party understands that the FAQ refers to the case where the

[15] Commission Decision 2000/520/EC of 26.7.2000 - O. J. L 215/7 of 25.8.2000

key coded data is sent to a recipient in the US (for instance, the pharmaceutical company), which receives only key-coded data and will never be aware of the identity of the patients which is known and will be known in case of need for treatment only to the medical professional/researcher in the EU, but never to the company in the US.

Anonymous data

"Anonymous data" in the sense of the Directive can be defined as any information relating to a natural person where the person cannot be identified, whether by the data controller or by any other person, *taking account of all the means likely reasonably to be used either by the controller or by any other person* to identify that individual. "Anonymised data" would therefore be anonymous data that previously referred to an identifiable person, but where that identification is no longer possible. Recital 26 also refers to this concept when it reads that "*the principles of protection shall not apply to data rendered anonymous in such a way that the data subject is no longer identifiable*". Again, the assessment of whether the data allow identification of an individual, and whether the information can be considered as anonymous or not depends on the circumstances, and a case-by-case analysis should be carried out with particular reference to the extent that the means are likely reasonably to be used for identification as described in Recital 26. This is particularly relevant in the case of statistical information, where despite the fact that the information may be presented as aggregated data, the original sample is not sufficiently large and other pieces of information may enable the identification of individuals.

Example No. 18: Statistical surveys and combination of scattered information

Apart from their general obligation to respect data protection rules, in order to ensure anonymity of the statistical surveys, statisticians are subjected to a specific duty of professional secrecy, and under those rules it is forbidden for them to publish non anonymous data. This obliges them to publish aggregated statistical data which cannot possibly be attributed to an identified person behind the statistics. This rule is particularly relevant concerning the publication of census data. In each situation a threshold should be determined under which it is deemed possible to identify the persons concerned. If a criterion appears to lead to identification in a given category of persons, however large (i.e. only one doctor operates in a town of 6000 inhabitants), this "discriminating" criterion should be dropped altogether or other criteria be added to "dilute" the results on a given person so as to allow for statistical secrecy.

Example No. 19: Publication of video surveillance

A shopkeeper installs a camera surveillance system in his shop. He publishes, in his shop, the pictures of thieves who have been caught by means of the camera surveillance system. After police intervention, he blanks out the faces of the thieves, by darkening them. However, even after this operation, there still exists a possibility that the persons on the photos can be recognized by their friends, relatives or neighbours, because of the fact that e.g. their figure, haircut and clothes are still recognizable.

4. FOURTH ELEMENT: "NATURAL PERSON"

The protection afforded by the rules of the Directive applies to natural persons, that is, to human beings. The right to the protection of personal data is, in that sense, a universal one that is not restricted to nationals or residents in a certain country. Recital

2 of the Directive explicitly makes this point by stating that "*data processing systems are designed to serve man*" and that they "*must, whatever the nationality or residence of natural persons, respect their fundamental rights and freedoms*".

The concept of natural person is referred to in Article 6 of the Universal Declaration of Human Rights, according to which "*Everyone has the right to recognition everywhere as a person before the law*". Member States' legislation, usually in the field of Civil Law, outlines more precisely the concept of personality of human beings, understood as the capacity to be the subject of legal relations, starting with the birth of the individual and ending with his death. Personal data are therefore data relating to identified or identifiable *living individuals* in principle. This raises a number of questions for the purposes of this analysis.

Data on dead persons

Information relating to dead individuals is therefore in principle not to be considered as personal data subject to the rules of the Directive, as the dead are no longer natural persons in civil law. However, the data of the deceased may still indirectly receive some protection in certain cases.

On the one hand, the data controller may not be in a position to ascertain whether the person to whom the data relate is still living or may be dead. Or even if he may do so, the information on the dead may be processed under the same regime as that on the living without distinction. As the data controller is subject to the data protection obligations imposed by the Directive as regards the data on living individuals, it will probably be easier for him in practice to process also the data on the dead in the way imposed by the data protection rules, rather than to separate the two sets of data.

On the other hand, the information on dead individuals may also refer to living persons. For instance, the information that the dead Gaia suffered from haemophilia indicates that her son Titius also suffers from the same disease, as it is linked to a gene contained in the X-chromosome. Thus, where the information which is data on the dead can be considered to relate at the same time also to the living and be personal data subject to the Directive, the personal data of the deceased may indirectly enjoy the protection of data protection rules.

Thirdly, information on deceased persons may be subject to specific protection granted by sets of rules other than data protection legislation, drawing the lines of what some call "*personalitas praeterita*". The obligation of confidentiality of medical staff does not end with the death of the patient. National legislation on the right to one's own image and honour may grant also protection to the memory of the dead.

And fourthly, nothing prevents a Member State from extending the scope of the national legislation implementing the provisions of Directive 95/46/EC to areas not included in the scope thereof provided that no other provision of Community law precludes it, as the ECJ has recalled[16]. It is possible that some national legislator may decide to extend the provisions of national data protection law to some aspects

16 Judgment of the European Court of Justice C-101/2001 of 06/11/2003 (Lindqvist), § 98

concerning processing data on deceased persons, where a legitimate interest may justify it[17].

Unborn children

The extent to which data protection rules may apply before birth depends on the general position of national legal systems about the protection of unborn children. To take mainly account of inheritance rights, some Member States acknowledge the principle that children conceived but not yet born are considered as if they were born as far as benefits are concerned (and thus can receive a heritage or accept a donation), subject to the condition that they may effectively be born. In other Member States, specific protection is given by particular legal provisions, also subject to the same condition. To determine whether national data protection provisions protect also information on unborn children, that general approach of the national legal system should be considered, together with the idea that the purpose of data protection rules is to protect the individual.

A second question is posed by the consideration that the legal system's general response relies on the expectation that the situation of unborn children is limited in time to the period of pregnancy. It does not take account of the fact that this situation may actually last considerably longer, as in the case of frozen embryos. Finally, specific legal responses may be found in particular provisions on reproduction techniques, dealing with the use of medical or genetic information about embryos.

Legal persons

As the definition of personal data refers to individuals, i.e. natural persons, information relating to legal persons is in principle not covered by the Directive, and the protection granted by it does not apply[18]. However, certain data protection rules may still indirectly apply to information relating to businesses or to legal persons, in a number of circumstances.

Some provisions of the e-privacy Directive 2002/58/EC extend to legal persons. Article 1 thereof provides that "*2. The provisions of this Directive particularise and complement Directive 95/46/EC for the purposes mentioned in paragraph 1. Moreover, they provide for protection of the legitimate interests of subscribers who are legal persons.*" Accordingly, Articles 12 and 13 extend the application of some provisions concerning directories of subscribers and unsolicited communication also to legal persons.

Information about legal persons may also be considered as "relating to" natural persons on their own merits, in accordance with the criteria set out in this document. This may be the case where the name of the legal person derives from that of a natural person. Another case may be that of corporate e-mail, which is normally used by a certain employee, or that of information about a small business (legally speaking an "object" rather than a legal person), which may describe the behaviour of its owner. In all these cases, where the criteria of "content", "purpose" or "result" allow the information on

---

[17] Minutes of the Council of the European Union, 8.2.1995, document 4730/95: "Re Article 2(a) "*The Council and the Commission confirm that it is for the Member States to lay down whether and to what extent this Directive shall be applied to deceased persons.*"

[18] Recital 24 of the Directive: "*Whereas the legislation concerning the protection of legal persons with regard to the processing of data which concerns them is not affected by this Directive*;"

the legal person or on the business to be considered as "relating" to a natural person, it should be considered as personal data, and the data protection rules should apply.

The European Court of Justice has made clear that nothing prevents the Member States from extending the scope of the national legislation implementing the provisions of the Directive to areas not included within the scope thereof, provided that no other provision of community law precludes it[19]. Accordingly some Member States such as Italy, Austria or Luxembourg have extended the application of certain provisions of national law adopted pursuant to the Directive (such as those on security measures) to the processing of data on legal persons.

As in the case of information on dead people, practical arrangements by the data controller may also result in data on legal person being subject de facto to data protection rules. Where the data controller collects data on natural and legal persons indistinctly and includes them in the same sets of data, the design of the data processing mechanisms and the auditing system may be set up so as to comply with data protection rules. In fact, it may be easier for the controller to apply the data protection rules to all sorts of information in his files than to try to sort out what refers to natural and what to legal persons.

IV. WHAT HAPPENS IF THE DATA FALL OUTSIDE OF THE DEFINITION?

As we have seen throughout this document, in different circumstances information may be considered not to be personal data. This is the case where the data cannot be considered to relate to an individual, or because the individual cannot be considered to be identified or identifiable. When the information that is processed does not fall within the concept of "personal data", the consequence is that the Directive does not apply, pursuant to Article 3 thereof. This does not mean, though, that individuals may be deprived of any kind of protection in the particular situation. We should take into account the following considerations.

If the Directive does not apply, national data protection law may apply. As laid down in Article 34, the Directive is addressed to the Member States. Outside of its scope, Member States are not subject to the obligations it imposes, basically to bring into force the laws, regulations and administrative provisions necessary to comply with it. However, as the European Court of Justice has made clear, nothing prevents Member States from extending the scope of the national legislation implementing the provisions of the Directive to areas not included within the scope thereof, provided that no other provision of community law precludes it. It may therefore very well happen that certain situations not involving processing of personal data as defined in the Directive are nevertheless subject to protective measures under national law. This may for instance apply to a subject like key-coded data, regardless of whether it is personal data or not.

Where data protection rules do not apply, certain activities may still constitute an interference with Article 8 of the European Convention on Human Rights, which protects the right to private and family life, in the light of the far-reaching jurisprudence of the ECHR. Other sets of rules, such as torts law, criminal law or anti-discrimination law may also provide protection to individuals in those cases where data protection rules do not apply and various legitimate interests may be at stake.

---

[19] Judgment of the European Court of Justice C-101/2001 of 06.11.2003 (Lindqvist), § 98

V. CONCLUSIONS

In this opinion the Working Party has provided guidance on the way in which the concept of personal data in Directive 95/46/EC and related community legislation should be understood and how it should be applied in different situations.

As a general consideration it has been noted that the European lawmaker intended to adopt a broad notion of personal data, but this notion is not unlimited. It should always be kept in mind that the objective of the rules contained in the Directive is to protect the fundamental rights and freedoms of individuals, in particular their right to privacy, with regard to the processing of personal data. These rules were therefore designed to apply to situations where the rights of individuals could be at risk and hence in need of protection. The scope of the data protection rules should not be overstretched, but unduly restricting the concept of personal data should also be avoided. The Directive has defined its scope, excluding a number of activities, and allows flexibility in the application of rules to activities that are within its scope. Data protection authorities play an essential role in finding an appropriate balance in this application (see paragraph II).

The Working Party's analysis has been based on the four main "building blocks" that can be distinguished in the definition of "personal data": i.e. "any information", "relating to", "an identified or identifiable", "natural person". These elements are closely intertwined and feed on each other, but together determine whether a piece of information should be considered as "personal data". The analysis is supported by examples from the national practice of European DPAs.

- The first element – "any information" – calls for a wide interpretation of the concept, regardless of the nature or content of the information, and the technical format in which it is presented. This means that both objective and subjective information about a person in whatever capacity may be considered as "personal data", and irrespective of the technical medium on which it is contained. The opinion also discusses biometric data and the legal distinctions with human samples from which they may be extracted (see paragraph III.1).

- The second element – "relating to" – has so far been often overlooked, but plays a crucial role in determining the substantive scope of the concept, especially in relation to objects and new technologies. The opinion provides three alternative elements – i.e. content, purpose or result – to determine whether information "relates to" an individual. This also covers information that may have a clear impact on the way in which an individual is treated or evaluated (see paragraph III.2).

- The third element – "identified or identifiable" – focuses on the conditions under which an individual should be considered as "identifiable", and especially on "the means likely reasonably to be used" by the controller or by any other person to identify that person. The particular context and circumstances of a specific case play an important role in this analysis. The opinion also deals with "pseudonymised data" and the use of "key-coded data" in statistical or pharmaceutical research (see paragraph III.3).

- The fourth element – "natural person" – deals with the requirement that "personal data" are about "living individuals". The opinion also discusses the interfaces with data on deceased persons, unborn children and legal persons (see paragraph III.4).

The opinion finally discusses what happens if data fall outside the scope of the definition of "personal data". Different solutions may be available to deal with issues in these cases, including national legislation outside the scope of the Directive, provided that other community law is respected (see paragraph IV).

The Working Party invites all interested parties to carefully study the guidance provided in this opinion and to take it into account when interpreting and applying provisions of national law in line with Directive 95/46/EC.

The members of the Working Party, mostly representatives of supervisory data protection authorities at national level, are committed to further developing the guidance provided in this opinion within their own jurisdictions and to ensuring a proper application of their national law in line with Directive 95/46/EC.

The Working Party intends to apply and develop the guidance provided in this opinion wherever appropriate, and to carefully take it into account in its further work, particularly when dealing with topics such as Identity Management in the context of e-Government and e-Health, as well as in the RFID context. As to the latter subject, the Working Party intends to contribute to a further analysis of the way in which data protection rules may impact on the use of RFIDs and of the possible need for additional measures that may be necessary in order to ensure a proper respect of data protection rights and interests in that context.

The Working Party would finally also welcome any feedback from interested parties and supervisory authorities on their practical experience with the guidance provided in this opinion, including any additional examples to those mentioned in this document. It intends to revisit the subject in due course, with a view to further enhancing the common understanding of the key concept of personal data, and ensuring a harmonized application and a better implementation of Directive 95/46/EC and related community legislation on that basis.

For the Working Party

The Chairman
Peter SCHAAR